

```

1 S:\Python35\python.exe S:/pycharm/sufe_jianmo/data.py
2 <class 'pandas.core.frame.DataFrame'>
3 RangeIndex: 30000 entries, 0 to 29999
4 Data columns (total 11 columns):
5 REPORT_ID          30000 non-null int64
6 ID_CARD            30000 non-null object
7 LOAN_DATE          30000 non-null object
8 AGENT              8952 non-null object
9 IS_LOCAL           30000 non-null object
10 WORK_PROVINCE      27742 non-null float64
11 EDU_LEVEL          26942 non-null object
12 MARRY_STATUS       30000 non-null object
13 SALARY             8864 non-null float64
14 HAS_FUND           29998 non-null float64
15 Y                  30000 non-null int64
16 dtypes: float64(3), int64(2), object(6)
17 memory usage: 2.5+ MB
18 None
19          REPORT_ID  WORK_PROVINCE      SALARY
   HAS_FUND          Y
20 count  3.000000e+04  27742.000000  8864.000000  29998.
   000000  30000.000000
21 mean   3.258910e+06  308176.414137    3.618682    0.
   455364    0.062500
22 std    1.563711e+06  101940.879912    1.415137    0.
   498012    0.242065
23 min    8.787000e+03  110000.000000    1.000000    0.
   000000    0.000000
24 25%    2.021495e+06  230000.000000    3.000000    0.
   000000    0.000000
25 50%    3.436337e+06  330100.000000    3.000000    0.
   000000    0.000000

```

```

26 75%      4.634978e+06  370000.000000      4.000000      1.
    000000      0.000000
27 max      5.631528e+06  650000.000000      7.000000      1.
    000000      1.000000
28 本地与非本地因素中客户逾期分析
29 未逾期客户人数
30 本地籍      16275
31 非本地籍      11850
32 Name: IS_LOCAL, dtype: int64
33 逾期客户人数
34 本地籍      1231
35 非本地籍      644
36 Name: IS_LOCAL, dtype: int64
37 本地户籍中未逾期比例是 0.929681252142      逾期比例是: 0.
    0703187478579
38 非本地户籍中逾期比例是 0.948455258524      逾期比例是: 0.
    0515447414759
39 学历与是否是逾期客户的分析
40 未逾期客户人数:
41 专科      7694
42 本科      6222
43 高中      5713
44 专科及以下      3870
45 初中      1269
46 其他      284
47 硕士研究生      150
48 博士研究生      50
49 硕士及以上      33
50 Name: EDU_LEVEL, dtype: int64
51 逾期客户人数:
52 专科及以下      582
53 高中      349

```

```

54 本科          314
55 专科          303
56 初中          86
57 其他          17
58 硕士研究生      3
59 硕士及以上      2
60 博士研究生      1
61 Name: EDU_LEVEL, dtype: int64
62 专科 中的未逾期比例 0.962110791547    逾期比例是 0.
    0378892084532
63 本科 中的未逾期比例 0.951958384333    逾期比例是 0.
    0480416156671
64 高中 中的未逾期比例 0.942428241504    逾期比例是 0.
    0575717584955
65 专科及以下 中的未逾期比例 0.869272237197    逾期比例是 0
    .130727762803
66 初中 中的未逾期比例 0.936531365314    逾期比例是 0.
    0634686346863
67 其他 中的未逾期比例 0.943521594684    逾期比例是 0.
    0564784053156
68 硕士研究生 中的未逾期比例 0.980392156863    逾期比例是 0
    .0196078431373
69 博士研究生 中的未逾期比例 0.980392156863    逾期比例是 0
    .0196078431373
70 硕士及以上 中的未逾期比例 0.942857142857    逾期比例是 0
    .0571428571429
71 婚姻情况与逾期情况分析
72 未逾期客户人数:
73 已婚          18091
74 未婚          8892
75 离异          676
76 离婚          270

```

```

77 其他          157
78 丧偶          39
79 Name: MARRY_STATUS, dtype: int64
80 逾期客户人数:
81 已婚          1108
82 未婚          674
83 离婚          44
84 离异          43
85 其他          4
86 丧偶          2
87 Name: MARRY_STATUS, dtype: int64
88 丧偶 中的未逾期比例 0.951219512195    逾期比例是 0.
    0487804878049
89 其他 中的未逾期比例 0.975155279503    逾期比例是 0.
    0248447204969
90 已婚 中的未逾期比例 0.942288660868    逾期比例是 0.
    0577113391322
91 未婚 中的未逾期比例 0.929542128371    逾期比例是 0.
    0704578716287
92 离婚 中的未逾期比例 0.859872611465    逾期比例是 0.
    140127388535
93 离异 中的未逾期比例 0.940194714882    逾期比例是 0.
    0598052851182
94 公积金（0-无公积金, 1-有公积金）与逾期客户分析
95 0 中的未逾期比例 0.935855061819    逾期比例是 0.
    0641449381809
96 1 中的未逾期比例 0.939458272328    逾期比例是 0.
    060541727672
97 salary的基本情况
98 count          8864.000000
99 mean           3.618682
100 std           1.415137

```

```
101 min          1.000000
102 25%          3.000000
103 50%          3.000000
104 75%          4.000000
105 max          7.000000
106 Name: SALARY, dtype: float64
107 有工资记录下的逾期比例分析
108 1 中的未逾期比例 0.742268041237 逾期比例是 0.
    257731958763
109 2 中的未逾期比例 0.836353340883 逾期比例是 0.
    163646659117
110 3 中的未逾期比例 0.87602224403 逾期比例是 0.
    12397775597
111 4 中的未逾期比例 0.930562116202 逾期比例是 0.
    0694378837978
112 5 中的未逾期比例 0.94422310757 逾期比例是 0.
    0557768924303
113 6 中的未逾期比例 0.951578947368 逾期比例是 0.
    0484210526316
114 7 中的未逾期比例 0.958263772955 逾期比例是 0.
    0417362270451
115
```